

Colegio España
Concepción 2020
Profesora Karinna Linares Avello



Guía de Aprendizaje N°6
Medidas de Dispersión
Tercero Medio Electivo

Nombre:	Curso:	Fecha:

Objetivo de Aprendizaje: Resolver problemas que involucren los conceptos de media muestral, desviación estándar, varianza, coeficiente de variación y correlación muestral entre dos variables, tanto de forma manuscrita como haciendo uso de herramientas tecnológicas digitales.

Importante: No es obligación imprimir esta guía, puedes copiarla y desarrollarla en tu cuaderno, estudiarla desde tu computador o dispositivo móvil. Consultas al correo electrónico karinna@cesp.cl

MEDIDAS DE DISPERSION PARA DATOS AGRUPADOS

Las medidas de dispersión son parámetros estadísticos que indican cuánto se alejan los datos respecto de la media aritmética. Es decir, indican la variabilidad de los datos. Las medidas de dispersión más utilizados son el rango, la desviación media y la desviación estándar o típica.

DESVIACIÓN ESTÁNDAR

La desviación estándar mide el grado de dispersión de los datos con respecto a la media. Se denota como σ para la población, o bien s para una muestra.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot f_i}{n}}$$

Donde:

x_i: marca de clase
$\bar{x}$: media aritmética
f_i: frecuencia absoluta de cada intervalo
n: número total de datos

Mientras menor sea la desviación estándar, los datos son más homogéneos, es decir, a menor dispersión mayor homogeneidad, y viceversa.

VARIANZA

Es otro parámetro utilizado para medir la dispersión de los valores de una variable respecto a la media. Corresponde a la media aritmética de los cuadrados de las desviaciones respecto de la media.

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot f_i}{n}$$

Donde:

x_i: marca de clase
$\bar{x}$: media aritmética
f_i: frecuencia absoluta de cada intervalo
n: número total de datos

Dado que la varianza corresponde al cuadrado de la desviación estándar, está expresada en unidades cuadradas.

COEFICIENTE DE VARIACIÓN

Permite determinar la razón existente entre la desviación estándar (σ) y la media ($\bar{x}$).

$$CV = \frac{\sigma}{\bar{x}}$$

El CV no tiene unidades de medida, por lo que permite la comparación de variables sin importar sus magnitudes ni lo que representan.

EJEMPLO N°1: DESVIACIÓN ESTÁNDAR Y VARIANZA PARA DATOS AGRUPADOS

Dada la siguiente tabla de frecuencias: Calcular la desviación estándar, la varianza y el coeficiente de variación.

Intervalo	x_i	f_i	$x_i \cdot f_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
[10, 20)	15	1	15	-28,33	802,59	802,59
[20, 30)	25	8	200	-18,33	335,99	2.687,92
[30, 40)	35	10	350	-8,33	69,39	693,9
[40, 50)	45	9	405	1,67	2,79	25,11
[50, 60)	55	8	440	11,67	136,19	1.089,52
[60, 70)	65	4	260	21,67	469,59	1.878,36
[70, 80)	75	2	150	31,67	1.002,99	2.005,98
		n = 42	1.820			9.183,38

Para calcular la desviación estándar, se requiere obtener la media aritmética:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i \cdot f_i}{n} = \frac{1.820}{42} = 43{,}33$$

La desviación estándar está dada por la expresión:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot f_i}{n}} = \sqrt{\frac{9.183{,}38}{42}} = \sqrt{218{,}65} = 14{,}79$$

La varianza está dada por la expresión:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot f_i}{n} = \frac{9.183{,}38}{42} = 218{,}65$$

El coeficiente de variación está dado por la expresión:

$$CV = \frac{\sigma}{\bar{x}} = \frac{14{,}79}{43{,}33} = 0{,}3413$$

EJEMPLO N°2: DESVIACIÓN ESTÁNDAR Y VARIANZA PARA DATOS AGRUPADOS

Dada la siguiente tabla de frecuencias: Calcular la desviación estándar y la varianza.

Intervalo	x_i	f_i	$x_i \cdot f_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
[10, 15)	12,5	3	37,5	-9,29	86,30	258,9
[15, 20)	17,5	5	87,5	-4,29	18,40	92
[20, 25)	22,5	7	157,5	0,71	0,50	3,5
[25, 30)	27,5	4	110	5,71	32,60	130,4
[30, 35)	32,5	2	65	10,71	114,70	229,4
		n = 21	457,5			714,2

Para calcular la desviación estándar, se requiere obtener la media aritmética:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i \cdot f_i}{n} = \frac{457{,}5}{21} = 21{,}79$$

La desviación estándar está dada por la expresión:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot f_i}{n}} = \sqrt{\frac{714{,}2}{21}} = \sqrt{34{,}00} = 5{,}83$$

La varianza está dada por la expresión:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot f_i}{n} = \frac{714,2}{21} = 34,00$$

El coeficiente de variación está dado por la expresión:

$$CV = \frac{\sigma}{\bar{x}} = \frac{5,83}{21,79} = 0,2676$$

CORRELACIÓN

La correlación indica el grado de asociación de dos variables, la influencia que pueda tener una sobre la otra, lo que a veces permite encontrar funciones que predicen ciertos comportamientos.

La correlación se mide usando el coeficiente de correlación lineal de Pearson (r). Este coeficiente fluctúa en el intervalo [-1, 1].

$$r = \frac{cov(x,y)}{s_x \cdot s_y}$$

$$cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n}$$

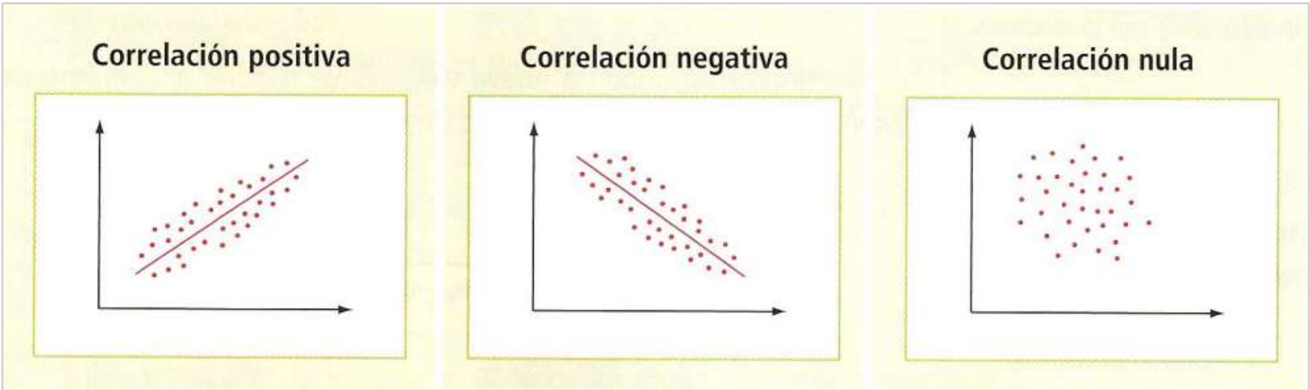
$cov(x,y) = covarianza$

$s_x = desviación\ estándar\ de\ x$

$s_y = desviación\ estándar\ de\ y$

Análisis del coeficiente de correlación:

- Si r es positivo, la relación lineal entre las variables es directa. Se dice que la correlación es positiva.
- Si r es negativo, la relación lineal entre las variables es inversa. Se dice que la correlación es negativa.
- Si r = 0, no existe relación lineal entre las variables. Se dice que la correlación es nula.
- Si r = 1, existe una relación de dependencia total directa entre las variables. Es decir, si una de ellas aumenta (o disminuye), la otra aumenta (o disminuye) en igual proporción.
- Si r = -1, existe una relación de dependencia total inversa entre las variables. Es decir, si una de ellas aumenta (o disminuye), la otra disminuye (o aumenta) en igual proporción.



EJEMPLO N°1: CORRELACIÓN - COEFICIENTE DE PEARSON

X	Y	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
2	4	-2	-2	4	4	4
3	5	-1	-1	1	1	1
4	6	0	0	0	0	0
5	7	1	1	1	1	1
6	8	2	2	4	4	4
				10	10	10

- Media Aritmética para variable x:

$$\bar{x} = \frac{2+3+4+5+6}{5} = \frac{20}{5} = 4$$

- Media Aritmética para variable y:

$$\bar{y} = \frac{4+5+6+7+8}{5} = \frac{30}{5} = 6$$

- Covarianza:

$$cov(x, y) = \frac{10}{5} = 2$$

- Desviación estándar variable x:

$$s_x = \sqrt{\frac{10}{5}} = \sqrt{2} = 1{,}41 \ldots$$

- Desviación estándar variable y:

$$s_y = \sqrt{\frac{10}{5}} = \sqrt{2} = 1{,}41 \ldots$$

- Coeficiente de Pearson:

$$r = \frac{2}{\sqrt{2} \cdot \sqrt{2}} = \frac{2}{2} = 1$$

∴ Existe una relación de dependencia total directa entre las variables.

EJEMPLO N°2: CORRELACIÓN - COEFICIENTE DE PEARSON

X	Y	$(x_i - \overline{x})$	$(y_i - \overline{y})$	$(x_i - \overline{x})(y_i - \overline{y})$	$(x_i - \overline{x})^2$	$(y_i - \overline{y})^2$
2	7	-2	1	-2	4	1
3	8	-1	2	-2	1	4
4	6	0	0	0	0	0
5	4	1	-2	-2	1	4
6	5	2	-1	-2	4	1
				-8	10	10

- Media Aritmética para variable x:

$$\bar{x} = \frac{2 + 3 + 4 + 5 + 6}{5} = \frac{20}{5} = 4$$

- Media Aritmética para variable y:

$$\bar{y} = \frac{7 + 8 + 6 + 4 + 5}{5} = \frac{30}{5} = 6$$

- Covarianza:

$$cov(x, y) = \frac{-8}{5} = -1{,}6$$

- Desviación estándar variable x:

$$s_x = \sqrt{\frac{10}{5}} = \sqrt{2} = 1{,}41 \ldots$$

- Desviación estándar variable y:

$$s_y = \sqrt{\frac{10}{5}} = \sqrt{2} = 1{,}41 \ldots$$

- Coeficiente de Pearson:

$$r = \frac{-1{,}6}{\sqrt{2} \cdot \sqrt{2}} = \frac{-1{,}6}{2} = -0{,}8$$

∴ La correlación es negativa, es decir, la relación lineal entre las variables es inversa.

EJERCICIOS

1. En la siguiente tabla de se muestran los años de servicio de una muestra de 100 empleados de un banco. Completa la tabla como en los ejemplos de la guía. Luego, calcula la desviación estándar y la varianza.

Años	N° Empleados
0 – 2	40
3 – 5	25
6 – 8	20
9 – 11	10
12 – 14	5

2. La siguiente tabla de distribución muestra las alturas de un curso de educación media de cierto colegio. Completa la tabla como en los ejemplos de la guía. Luego, calcula la desviación estándar y la varianza.

Alturas	N° de alumnos
[150, 155)	3
[155, 160)	7
[160, 165)	6
[165, 170)	4
[170, 175)	5

3. La tabla de distribución de frecuencias muestra la puntuación obtenida por 1.800 alumnos de 5° a 8° básico en un cuestionario de cultura general. Completa la tabla como en los ejemplos de la guía. Luego, calcula la desviación estándar y la varianza.

Puntaje	Frecuencia
0 – 2	21
3 – 5	50
6 – 8	110
9 – 11	241
12 – 14	423
15 – 17	457
18 – 20	275
21 – 23	134
24 – 26	66
27 – 29	23

4. Calcular la correlación de Pearson para el siguiente conjunto de datos.

X	10	8	13	9	11	14	6	4	12	7	5
Y	9,14	8,14	8,74	8,77	9,26	8,10	6,13	3,10	9,13	7,26	4,74